

AD \_\_\_\_\_

AWARD NUMBER:

Award Number W81XWH-12-1-0123

TITLE:

Delineation of Methyl-DNA Binding Protein Interactions in the Prostate Cancer Genome (PC110091)

PRINCIPAL INVESTIGATOR:

Roderick T Hori, PhD

CONTRACTING ORGANIZATION:

University of Tennessee Health Science Center

Memphis, TN 38163 - 0001

REPORT DATE:

March 2014

TYPE OF REPORT:

Final Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE March 2014		2. REPORT TYPE Final Report		3. DATES COVERED 1 July 2012 – 31 December 2013	
4. TITLE AND SUBTITLE Delineation of Methyl-DNA Binding Protein Interactions in the Prostate Cancer Genome (PC110091)				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-12-1-0123	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Roderick T Hori, PhD  E-Mail: rhori@uthsc.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)  University of Tennessee Health Science Center Dept: Finance and Operations 62 S. Dunlap Street, Suite 300 Memphis, TN 38163-0001				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES Prostate Cancer, Methylated DNA, Methyl-CpG Binding Domain, Chromatin Immunoprecipitation					
14. ABSTRACT The purpose of this study is to generate a genome-wide association profile of Methyl-CpG Domain-containing (MBD) proteins, such as MeCP2, MBD1, MBD2 and MBD4, in malignant prostate cancer cells and matched normal or benign prostate cells using Chromatin Immunoprecipitation followed by Next Generation Sequencing (ChIP-Seq). ChIP-Seq experiments to examine those four proteins have been performed on two or three matched sets of prostate tissues. The results have been analyzed using a combination of software packages including Genomics Suite (Partek), RSEG and Galaxy to identify sets of common association regions or sites for each protein. The genes that overlap these common sites were identified and the genes among these with increased or decreased gene expression by microarray analysis have been listed. In parallel, RNA expression profiles from the same tissues were generated to allow comparison of differential patterns of gene expression with differential patterns of MBD protein association. Microarray analysis has been performed and has identified genes that are up-regulated and down-regulated in Stage 3 prostate cancer cells. This data has been useful in facilitating the bioinformatics analysis of the ChIP-Seq results and is also useful on its own to identify genes with altered patterns of regulation in malignant tissues.					
15. SUBJECT TERMS nothing listed					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	36	19b. TELEPHONE NUMBER (include area code)

## Table of Contents

	<u>Page</u>
<b>Introduction.....</b>	<b>4</b>
<b>Body.....</b>	<b>4</b>
<b>Key Research Accomplishments.....</b>	<b>13</b>
<b>Reportable Outcomes.....</b>	<b>15</b>
<b>Conclusion.....</b>	<b>16</b>
<b>References.....</b>	<b>17</b>
<b>Appendices.....</b>	<b>N/A</b>
<b>Supporting Data.....</b>	<b>18</b>

## **Introduction**

Tumors exhibit abnormal epigenetic regulation (1). A key epigenetic event is DNA methylation, typically at the dinucleotide CpG. Both hypomethylation and hypermethylation of DNA is observed in cancer cells. Generally, the former leads to activation of protooncogenes and the latter to repression of tumor suppressors during the initiation and progression of cancer. The Methyl-CpG-Binding Domain (MBD) family of proteins is among the ones that bind methylated DNA directly and this association is critical in regulating the outcome of DNA methylation (2,3). The MBD family includes proteins named MBD1, MBD2, MBD4 and MeCP2. (MBD3 contains the MBD but does not bind methylated DNA directly.) The molecular events leading from DNA methylation or demethylation to gene regulation during cancer progression are poorly defined. In this study, the binding of MBD proteins throughout the entire genome of several malignant prostate tumors and matched normal or benign tissues will be identified. This provides novel fundamental information about epigenetic regulation during carcinogenesis. Also, the differences in these patterns will identify new genes regulated by MBD binding and DNA methylation during prostate cancer progression. These genes potentially identify novel biomarkers or therapeutic targets. The identity of bound MBD protein will facilitate subsequent studies that examine the “downstream” molecular events at a specific gene during progression of prostate cancer.

## **Body**

As a brief overall summary, the genomic profiles of MBD proteins in malignant and matched normal or benign prostate tissue was examined by Chromatin Immunoprecipitation (ChIP) followed by Next Generation Sequencing (NGS or Seq). The details of the ChIP-Seq experiments are described below. The subsequent bioinformatics analysis of this project was performed to identify genes potentially targeted and regulated by MBD proteins – described below. This downstream informatics analysis was demanding because MBD proteins are associated with the genome over wide regions, a major difference from “typical” DNA-binding proteins such as CREB or C/EBP which bind relatively small and discrete recognition sites

and the bioinformatic analysis required increased effort. The pattern of MBD binding is similar to that observed with many histone modifications, another form of epigenetic regulation. One thing I have learned from this study is that additional information can be obtained and questions addressed with continued bioinformatic approaches, almost ad infinitum, so this analysis will be a continuing process. In the following summary, I have attempted to include representative and important analyses.

The first task includes the Chromatin Immunoprecipitations (ChIP) and Next Generation Sequencing (NGS or Seq) experiments. Five sets of matched malignant prostate carcinomas and normal/benign prostate tissues were obtained from the Cooperative Human Tissue Network (<http://www.chtn.nci.nih.gov>). These tissues were divided into two portions. The larger portion was prepared for ChIP experiments using the truChIP Tissue Chromatin Shearing Kit (Covaris, Woburn, MA). (The smaller portion of tissue was used for isolation of RNA - see below.) Briefly, the tissues were minced, treated with formaldehyde and this reaction was quenched per manufacturer's suggested protocol. Each tissue was disrupted and the separated cells that resulted were lysed. The nuclei were collected by centrifugation, then sonicated to lyse the nuclei and shear the chromatin in an Adaptive Focused Acoustics sonicator (Covaris). Chromatin immunoprecipitations were carried out using antibodies that recognize MeCP2, MBD1, MBD2 and MBD4 and the ChIP-IT High Sensitivity Kit (Active Motif, Carlsbad, CA). The resulting "ChIP DNA" was quantified using a Qubit fluorometer (Life Technologies, Grand Island, NY), then concentrated and purified by ethanol precipitation using linear acrylamide as a carrier and quantified again at the end using the Qubit fluorometer. The amounts of DNA obtained after the ChIP and ethanol precipitation made the next step, the preparation of libraries for Next-Generation Sequencing, technically challenging and required more time than allotted in the "Statement of Work". This delay led to the request and approval of a no-cost extension on the grant.

The SeqPlex DNA Amplification Kit (Sigma, St Louis, MO) was used to generate the libraries for NGS following the manufacturer's suggested protocol. ChIP

DNA from three matched sets of malignant prostate cancer and normal/benign prostate tissue were sequenced using the Ion Torrent (Life Technologies). Approximately 1 million sequencing “reads” were performed for each sample. With additional resources, a higher number of reads per sample might increase the number of identifiable peaks (that reach the appropriate statistical threshold).

The second task will be brought into the discussion at this point, prior to the bioinformatics analysis portion of Task 1, because the information from the microarrays performed in the second task was used in the bioinformatic analysis. In brief, RNA was isolated from the prostate tissues, a microarray was performed to identify genes that are differentially expressed in malignant prostate tissues and in the bioinformatic analysis, this information was correlated with binding of MBD proteins. An mRNA fraction was isolated from each tissue using an RNeasy Kit and a Qiacube (Qiagen USA, Valencia, CA). (In addition, microRNA fractions have been isolated from two sets of matched tissues to expand this study to include microRNAs in the future.) Equal amounts of mRNA from three normal prostate tissues and from three stage 3 malignant prostate carcinomas, respectively, were mixed and assayed using an Affymetrix microarray. Using this pooled set of mRNAs, the microarray experiment was used to generate a list of genes that are increased >2-fold and decreased >2-fold in the stage 3 prostate cancer tissue relative to normal tissue (see Tables 1 and 2 in Supporting Data for 98 most up-regulated and down-regulated genes). Genes of interest, namely their expression levels, will need to be validated by reverse transcriptase-quantitative PCR (RT-qPCR) using mRNAs from matched sets of tissues. The threshold of 2-fold is arbitrary and gene lists using different thresholds can be generated easily.

The bioinformatics analysis (part of Task 1) was used to address different questions. As a brief introduction of the bioinformatic analysis, “alignment” refers to the process of identifying (or aligning) the sequence reads against the reference genome, in this case the hg19 compilation of the human genome. This process records the sequencing reads attributable to each genomic region. After the

alignments are performed, they are compared between MBD-family ChIPs and either input DNA or an IgG-control ChIP to identify “peaks” or “regions” of aligned sequence reads within the genome. The chromosomal locations and corresponding genes within those peak regions are analyzed to identify potential MBD targets.

In one analysis, the alignments produced by the software of the Ion Torrent sequencer were used. These results were then processed (in collaboration with H. Chen) in a software package called RSEG (4) to identify the boundaries of the peaks from the ChIP-Seq experiments. RSEG was developed to identify the boundaries of broad, less discrete peaks such as those associated with histone modification. This property makes it useful for the analysis of MBD binding profiles. The peaks identified by RSEG were analyzed using the tools within the Galaxy website (<https://usegalaxy.org>) to identify genes that overlap the peaks. An alternative analysis is using the commercially available Genomics Suite – ChIP-Seq workflow (Partek, St Louis MO), which has been used in other publications (4, 5). The analysis using RSEG/Galaxy and Genomics Suite (Partek) will be described below.

The bioinformatics is an on-going task. The analysis above is being continued to extract additional information and results for publication. The bioinformatics analysis (in collaboration with Q Tran) is also being performed in a third way for comparison using Bowtie ([www.bowtie-bio.sourceforge.net](http://www.bowtie-bio.sourceforge.net)) to perform the alignments and MACS (<http://liulab.dfci.harvard.edu/MACS/>) to identify the peaks. This is ongoing and not described below.

The analysis of the MBD1 ChIP-Seq experiments using RSEG and Galaxy to identify and analyze **\*\*malignant-specific\*\*** MBD1 ChIP-Seq peaks is summarized next. MBD1 ChIP-SEQ was performed using three matched sets of malignant and normal/benign prostate tissues. As described above, RSEG was used to identify the MBD1 ChIP-Seq peaks from matched malignant and normal or benign tissues. Using Galaxy, the MBD1 ChIP-Seq peaks specific to the malignant tissue of each sample were identified by subtracting the peaks from the normal or benign tissue from the peaks present in the matched malignant prostate tissue. These resulting peaks from

each of three sets of samples were compared and 32 malignant-specific MBD1 peaks common to all three samples were identified (see Table 3). There is a large difference in the number of common peaks identified when all three samples are compared versus the pair-wise comparisons. When the malignant-specific MBD1 peaks were compared in a pair-wise manner, 164, 94 and 483 common peaks (or genomic regions) were identified (list of peaks not shown). The malignant-specific MBD1 peaks common to all three samples were analyzed to identify regions containing known genes. Of these 32 peaks, most are gene-containing region and this analysis yielded 31 malignant-specific peaks (see Table 4). These 31 gene-containing regions were examined using Integrated Genome Browser (IGB; [www.bioviz.org](http://www.bioviz.org)). Twenty-six of these regions contain 6 genes or less (data not shown). The relative expression level of the genes from these 26 regions in malignant versus normal prostate tissues was compared using the complete microarray results described above. Of these 26 regions, nineteen contain just one or two genes. (Note: One, some genes within these 26 regions were not found within the microarray and levels could not be determined from that experiment. Two, the regions beyond the 26 above were not investigated further because they contain between 10 and 33 genes. The large number of genes in these regions greatly diminishes the value of the analysis.) Of the genes overlapping the malignant-specific MBD1 peaks common to all three tissue samples, five genes had decreased levels of RNA in malignant tissues: SLC22A15, FAM110C, PLEKHG4B, PDCD6 and HRSP12 (listed in Table 4). Decreased expression was defined as RNA level less than 80% of that observed in normal tissues. Of these investigated genes, six genes had at least 1.4-fold increased RNA level in malignant relative to normal tissues: CASQ2, DPF3, KIAA0226, CCDC127, PTPRN2 and snoRNA72 (listed Table 4). As a reference point, DHFR RNA level was increased by 1.2-fold in the microarray. Some of the noteworthy features of these genes include the role of: FAM110C in the Akt signaling pathway and prostatitis, PLEKHG4 in phospholipid binding and Rho guanyl-nucleotide exchange factor activity and PDCD in programmed cell death, DPF3 as a subunit of BAF - a chromatin remodeling complex, KIAA0226's homology to PLEKHM1L, a plekstrin-homology domain contain protein and PTPRN2 as a receptor



with tyrosine phosphatase activity. Additional literature research investigating protein function and activity is ongoing. Additional experiments to evaluate their expression level in the matched tissues will be performed.

Many questions can be asked with this ChIP-Seq data. For example, the genomic regions containing the MBD1 peaks in each INDIVIDUAL tissue (as opposed to the analysis above examining what is common to all three sets) can be analyzed for the genes present and expression levels of those genes. As another example, the peaks common to only two tissues can be obtained by carrying out the pair-wise comparison of ChIP-Seq peaks. Based on the numbers listed above, this is less stringent but could identify genes that are important to the process of prostate cancer.

A similar analysis using RSEG and Galaxy was used to examine the MeCP2 ChIP-Seq experiments and identify and analyze malignant-specific MeCP2 ChIP-Seq peaks. MeCP2 ChIP-Seq was performed using the same three matched sets of malignant and normal/benign prostate tissues as used in the MBD1 study above. RSEG was used to identify the boundaries of the ChIP-Seq peaks. Using Galaxy, the malignant-specific MeCP2 ChIP-Seq peaks for each sample were identified as described above. These resulting peaks from each of three sets of samples were compared and 38 malignant-specific MeCP2 peaks common to all three samples were identified (see Table 5). For comparison, when the malignant-specific MeCP2 peaks were compared in a pair-wise manner between samples, there were 1835, 462 and 1815 common peaks were identified (peak lists not shown). The malignant-specific MeCP2 peaks common to all three matched samples were analyzed to identify regions that contained known genes within these common peaks. Of these 38 peaks, most contained genes and this analysis yielded 29 malignant-specific peaks (see Table 6). These 29 genomic regions were examined using Integrated Genome Browser. Twenty-three of these regions contain 7 genes or less (data not shown). The relative expression level of these genes in malignant versus normal prostate tissues was compared using the complete microarray results described

above. (NOTE: Of these 23 regions, fifteen contain just one or two genes. The remaining regions, which contain between 11 and 20 genes, will not be analyzed further at this time because the large number of genes.) Of the genes overlapping the malignant-specific MeCP2 peaks common to all three tissue samples, five genes had decreased levels of RNA in malignant tissues: SORC3, C11orf45, ACOXL, GINS1 and GRIA1 (listed in Table 6). Decreased expression was defined as RNA level less than 80% that observed in normal tissues. Of these investigated genes, six genes had at least 1.4-fold increased RNA level in malignant relative to normal tissues: RPS6KC1, CDH11, PTPRM, LYPD1, NCKAP5 and PTPRT (listed in Table 6). Some of the noteworthy features of these genes include the role of: SORC3 in vacuole sorting, the association of ACOXL with leukemias, GIN, an replication initiation factor which contains GINS1 as a subunit, RPS6KC1 in regulating protein synthesis, CDH11's involvement in signaling as a cadherin, PTPRM and PTPRT which are both receptors containing tyrosine phosphatase activity and LYPD1 (also known as C4.4) in tumor metastasis. Additional experiments to examine gene expression levels and additional literature research investigating protein function are needed.

The MBD4 ChIP-Seq experiments were analyzed in order to identify the malignant-specific peaks. In this case, the ChIP-Seq data from just two of three matched tissue sets used in the MBD1 and MeCP2 studies above was obtained. RSEG and Galaxy were used as described above to identify the malignant-specific MBD4 ChIP-Seq peaks for each sample. This analysis for MBD4 will be described in a bit more detail than for MBD1 and MeCP2 to provide some insight into the number of peaks that were typically identified in the ChIP-Seq experiments. In one set of matched tissues, 1173 malignant-specific peaks were identified out of 1,936 peaks identified from the malignant tissue. In the second set of matched tissues, 760 malignant-specific peaks were identified out of 1,130 peaks. When the malignant-specific peaks are compared, 143 overlapping malignant-specific peaks between these two tissues were identified (see Table 7). The genes within these peaks will be identified and expression will be examined using the previously discussed microarray and additional experiments such as RT-qPCR or microarrays.

MBD2 ChIP-Seq data was also obtained for only two of three matched sets of tissues used in the MBD1 and MeCP2 studies above. The analysis was performed as described for MBD4. The overlapping malignant-specific peaks will be analyzed for overlapping genes and gene expression will be examined further.

The analysis performed using Genomics Suite is described next. The first part of the analysis is aimed at identifying the sites where the methyl CpG-binding domain proteins are associated. Briefly, the raw high-throughput sequencing data was aligned against a reference human genome (hg19) to identify peaks (also called “regions”). The alignment was performed in pairs - experimental ChIP DNA and input DNA from malignant or normal/benign samples. The peaks that passed the criterion of a false discovery rate (FDR) equaling 0.2 or less were identified, then the peaks found in the experimental ChIP sample or not and those found in the input DNA or not were sorted. The peaks that were found in the experimental ChIP were filtered to find the ones that are specific to the ChIP (that is, not found in the input). The outcome of these operations is to identify peaks (i.e., regions within the genome) associated with the corresponding methyl CpG-binding domain protein (MBD1, MBD2, MBD4 or MeCP2).

MBD1 ChIP-Seq results were obtained from three matched sets of samples – (Sample or Patient) #3, #4 and #6. Using the Partek Genomics Suite software, the normal tissue samples 1897, 1491 and 1257 peaks (also called “regions”), respectively, were identified in the normal tissues and 1826, 1814 and 1355 peaks, respectively, were identified in the malignant tissues. Of the 1897 identified peaks (or “regions”) found in the normal tissue of sample #3, 1234 peaks were identified in the MBD1 ChIP and 1329 peaks were identified in the input DNA. This distribution was roughly similar for all samples (malignant, normal or benign and each ChIP – MBD1, MBD2, MBD4 and MeCP2). In our continued analysis, we found that 217 of the 1897 peaks from Sample #3 normal tissue passed the false discovery rate (FDR). Of those 217 peaks, 174 were present in the MBD1 ChIP-Seq samples. Of those 174 peaks,

42 were specific to the MBD1 ChIP (that is, the peaks were NOT present in the Input DNA). These MBD1-specific peaks are listed in Table 8 and are predicted to correspond to genomic regions where MBD1 is associated in the normal tissue from Sample #3. Continuing with the Partek Genomics Suite, when it was used to identify genes within 10 kb of the peaks, those 42 peaks yielded a set of 41 Transcript IDs (see Table 9). It should be noted that many of the Transcript IDs correspond to the same gene (symbol). Based on these results, the peaks (Table 8) and gene list (Table 9) indicate the genomic locations where MBD1 was bound in normal prostate tissue from sample #3. Each ChIP-Seq data set generates similar sized arrays of results, whether malignant, normal or benign or an immunoprecipitation (ChIP) of MBD1, MBD2, MBD4 or MeCP2. The results for MBD1 are shown in Tables 8-15 including a list of normal tissue-specific peaks and malignant specific peaks for Sample (or Patient) #3 and closest genes to normal-specific peaks and malignant-specific peaks for all three Samples (Patient #3, #4 and #6). This set of tables is listed in the Supporting Data at the end, but it is also listed below to help clarify what is present.

Table 8: Patient 3 Normal MBD1 Peaks

Table 9: Patient 3 Normal MBD1 Gene List

Table 10: Patient 4 Normal MBD1 Gene List

Table 11: Patient 6 Normal MBD1 Gene List

Table 12: Patient 3 Malignant MBD1 Peaks

Table 13: Patient 3 Malignant MBD1 Gene List

Table 1410: Patient 4 Malignant MBD1 Gene List

Table 15: Patient 6 Malignant MBD1 Gene List

NOTE: Tables for MBD1-specific Normal and Malignant Peaks for Samples #4 and #6 and all of the MBD2, MBD4 and MeCP2 results using Genomics Suite (Partek) were not included in an attempt to minimize the number of tables in this document.

I need to address an important technical point regarding the analysis performed on Partek Genomics Suite. One, as this analysis was being performed, the small number of common sites was surprising. I followed up this point in two ways. I used the “genome browser” function which allows me to visually compare peaks in different samples and found that some common sites seemed to be present that were not being identified by the software package. A visual inspection is one method for identifying the common peaks. While the visual inspection of one chromosome for comparison is possible, each autosomal and sex chromosome is presented separately so ChIP-Seq of each tissue would require 24 visual inspections (22 autosomal, X and Y chromosomes). This is impractical. That is part of the reason that RSEG/Galaxy and Bowtie/MACS was/is being used for this analysis. Two, I attempted to identify common peaks in three related sets of results simultaneously. This software package is supposed to have the capability to perform this analysis, but the output in three columns was not making sense. After discussions with customer support at Partek, it is likely there is a “bug”.

The first task was broken down into four parts and the first three parts (perform ChIPs, prepare samples for NGS and perform NGS) have been completed. Two matched malignant/normal tissue samples and one matched malignant/benign sample was used in our ChIP-Seq studies. The bioinformatic analysis is the fourth task and some important and representative analysis has been presented in this report. The bioinformatics analysis will be continued to generate additional information. The second task is identifying both genes that are differentially expressed in malignant prostate tissues and corresponding gene promoters that differentially bound by MBD proteins in malignant prostate tissues. Microarray experiment has been performed and those results have been used to generate correlations between sites of MBD family association and differentially regulated genes in malignant tissues. The steps involved in the two tasks have been detailed above.

The third task is the writing and submission of a manuscript. The delay involved in preparing samples for NGS and the time-consuming nature of the data analysis has meant that this task has not been completed. The results described above will be used in the future for both a manuscript and in grant applications.

### **Key Research Accomplishments**

- Chromatin Immunoprecipitation using matched malignant and normal/benign prostate tissues were performed with antibodies directed against MeCP2, MBD1, MBD2 and MBD4 (methyl-CpG binding domain containing proteins)
- “Next Generation Sequencing of ChIP experiments described was performed.
- The bioinformatics analysis of NGS data has been undertaken. Bioinformatic analysis is still on-going. Differences in the MBD association profiles between matched tissues were observed. Genes overlapping the binding sites of MBD proteins were identified. The microarray analysis was used to determine which of the overlapping genes were up-regulated or down-regulated.
- mRNAs has been isolated from each tissue. Small RNAs (miRNAs being of interest) have been isolated from some matched tissues.
- The mRNA from Stage 3 prostate cancer tissue and normal tissue was pooled, respectively, and analyzed using Affymetrix ST Human Gene 2.0 microarrays. The analysis identified genes that are up-regulated and down-regulated in malignant tissues – see Supporting Data.

## **Reportable Outcomes**

1. The association site for MBD1, MBD2, MBD4 and MeCP2 in matched sets of malignant and normal or benign prostate tissues
2. Using information in point #1 to identify the common association site for MBD1 and MeCP2 among three matched sets of tissues
3. Identification of genes that overlap with the common association sites in point #2
4. Identification of genes from point #3 that are up-regulated or down-regulated by microarray analysis of prostate tissues
5. Using information in point #1 to identify common association site for MBD1 and MeCP2 among two matched sets of tissues (that is a pair-wise comparison instead of the three way)
6. Using information in point #1 to identify common association site for MBD2 and MBD4 among two matched sets of tissues
7. Microarray analysis of malignant and normal prostate tissue to identify genes that are consistently up-regulated or down-regulated.
8. Additional question are being addressed by continuing the bioinformatics analysis

## Conclusions

This study has identified the association sites of MBD family members in matched malignant and normal or benign prostate tissues through ChIP-Seq experiments and bioinformatics analysis, then used these results to identify genes that overlap these regions. These results are important because they provide a genome-wide entry to understanding one type of epigenetic regulation that is often involved in cancer. The implication is that some of the sites or regions of MBD1, MBD2, MBD4 and MeCP2 association that are specific to malignant tissues are important to tumor progression. From a technical stand-point, this study is useful in contributing to early, but growing, literature that ChIP-Seq can be used on clinical (patient) samples.

The results are summarized by a variety of tables – in the supporting data - that list the malignant-specific association sites for these MBD-family members and the regions containing overlapping genes and the microarray analysis. This information generates two types of scientific hypotheses. One, that the identified genes are involved in tumor progression and that the association of MBD family members regulates their expression in a manner that leads to cancer. Two, that the identified genes may not be involved in tumor progression but their expression changes may be useful as biomarkers. In the future, I will propose studies to characterize those hypotheses for selected genes. By increasing the understanding of gene regulation leading to malignant cancers, a critical mass will be reached where the disease can be addressed in a differential manner. Cancer is such a complicated disease that ideas of “one size fits all” are naïve. However, a more robust understanding of the basic science underlying the process will allow clinicians to determine the specific pathways have been corrupted, how the pathways intertwine and the stage of the process and thus tailor treatment accordingly.



## References

1. Rodríguez-Paredes M, and Esteller M. (2011). Nat Med. 17(3):330-9. *Cancer epigenetics reaches mainstream oncology.*
2. Bogdanovic O and Veenstra GJ (2009). Chromosoma. 118(5):549-565. *DNA methylation and methyl-CpG binding proteins: developmental requirements and function.*
3. Parry L and Clarke AR (2011). Genes Cancer. 2(6):618-630. *The Roles of the Methyl-CpG Binding Proteins in Cancer.*
4. Song Q and Smith AD (2011). **Bioinformatics** 27 (6): 870-871. *Identifying dispersed epigenomic domains from ChIP-Seq data.*
5. Lo R and Mathews J (2012). Toxic Sciences 130(2):349-361. *High-Resolution Genome-wide Mapping of AHR and ARNT Binding Sites by ChIP-Seq*
6. Hewitt SC, [Li L](#) [Grimm SA](#), [Chen Y](#), [Liu L](#), [Li Y](#), [Bushel PR](#), [Fargo D](#), and [Korach KS](#) (2012). **Mol Endocrinol**, 26(5): 887–898. *Research Resource: Whole-Genome Estrogen Receptor  $\alpha$  Binding in Mouse Uterine Tissue Revealed by ChIP-Seq*

## Bibliography and Meeting Abstracts

None at the current time. A manuscript, grant applications and meeting presentation is being planned now that the analysis is nearing completion.

## Personnel Supported by this Grant

1. Roderick T Hori, PhD. Associate Professor, University of Tennessee HSC
2. Joycelynn Butler, B.S., Research Associate, University of Tennessee HSC

## Appendices

No Appendices

## **Supporting Data**

Table 1. Genes that are up-regulated in Stage 3 prostate tumors (top 98)

Table 2. Genes that are down-regulated in Stage 3 prostate tumors (top 98)

Table 3. Malignant-Specific MBD1 Peaks Common to All Samples

Table 4. Gene-Containing Regions Overlapping Malignant-Specific MBD1 Peaks

Table 5. Malignant-Specific MeCP2 Peaks Common to All Samples

Table 6. Gene-Containing Regions Overlapping Malignant-Specific MeCP2 Peaks

Table 7. Malignant-Specific MBD4 Peaks Common to the Two Patient Samples

Table 8. MBD1-Specific Peaks in Patient #3 Normal Tissue

Table 9. Genes Nearest MBD1-Specific Peaks in Patient #3 Normal Tissue

Table 10. Genes Nearest MBD1-Specific Peaks in Patient #4 Normal Tissue

Table 11. Genes Nearest MBD1-Specific Peaks in Patient #11 Normal Tissue

Table 12. MBD1-Specific Peaks in Patient #3 Malignant Tissue

Table 13. Genes Nearest MBD1-Specific Peaks in Patient #3 Malignant Tissue

Table 14. Genes Nearest MBD1-Specific Peaks in Patient #4 Malignant Tissue

Table 15: Genes Nearest MBD1-Specific Peaks in Patient #5 Malignant Tissue

TABLE 1: Genes up-regulated in Stage 3 Prostate Cancer

Probe ID	Stage 3	Normal	GeneAccession	Gene Symbol	mRNA Accession	Fold Increase
16888999	582.4666	13.2632	---	---	TCONS_I2_000:	43.91596001
16797447	101.4443	2.81261	BC073766	IGHV4-31	BC073766	36.06769944
16900148	291.9534	9.17729	ENST0000048481	IGKV2-24	ENST00000484	31.81258351
16734799	559.6259	17.6337	---	---	ENST00000357	31.73610461
16882744	226.7969	8.43171	ENST0000045318	IGKV1D-27	ENST00000453	26.89808428
17092187	1071.185	42.6476	NM_006911	RLN1	NM_006911	25.11712854
16877453	139.4926	6.79325	---	---	ENST00000421	20.53399462
16927869	253.0766	15.1123	ENST0000039033	IGLC7 // IGL	ENST00000390	16.7463543
16994426	43.19071	2.90542	---	---	---	14.8655804
16906620	934.7865	64.4569	NM_016192	TMEFF2	NM_016192	14.50250027
17086067	713.7975	50.5516	NR_015342	PCA3	NR_015342	14.12016504
16797494	113.5486	8.53564	ENST0000039060	IGHV3-23 //	ENST00000390	13.30288686
16987773	59.76891	4.57236	---	---	---	13.07178569
16734793	736.5161	61.2429	NM_030774	OR51E2	NM_030774	12.0261526
16900152	68.56734	5.88263	---	---	AY510106	11.65590277
16994424	45.59653	3.9256	---	---	---	11.61516587
16713395	44.09912	3.8074	---	---	---	11.58247319
16870453	400.378	36.6231	NM_000095	COMP	NM_000095	10.93240459
16721212	70.11575	6.46355	NM_001004753	OR51F2	NM_001004753	10.8478683
16797415	1435.142	136.309	BC092449	IGHA1	BC092449	10.52856223
16797583	53.47449	5.19404	AF067420	IGHA1	AF067420	10.29535187
16882750	203.5067	20.8347	ENST0000046269	IGKV2D-24	ENST00000462	9.767671464
16986655	295.422	30.5889	NM_003248	THBS4	NM_003248	9.657817051
16797512	172.9989	18.1594	ENST0000043814	IGHV4-31	ENST00000438	9.526706888
16797587	41.84327	4.42816	AK097017 // AK	IGHA1 // IGH	AK097017	9.449366734
16797389	807.075	85.6432	---	---	AY172958	9.42368818
16927853	1388.054	147.93	---	---	AY172962	9.383181234
16927834	698.1865	75.5693	---	---	ENST00000385	9.239019322
16818443	479.6125	51.9562	---	---	TCONS_I2_000:	9.231098058
16900098	561.5751	60.9755	BC016380	IGK@	BC016380	9.209851239
17095887	141.7376	15.7169	NM_017680	ASPN	NM_017680	9.018142208
16721210	54.4311	6.12543	---	---	ENST00000365	8.886080562
16900104	189.546	21.9098	ENST0000046416	IGKV1-6 // I	ENST00000464	8.651209027
16923068	253.2604	29.5595	---	---	TCONS_000290	8.567817453
16797371	313.0038	37.6674	AK125238	IGHA1	AK125238	8.309666999
17000168	664.37	80.7914	NM_004887	CXCL14	NM_004887	8.223274204
16797516	768.5807	96.3621	Y14737 // Y1473	IGHM // IGH	Y14737	7.975961306
16927858	1583.615	199.798	BC030984	IGLV1-44	BC030984	7.926100176
16976644	277.1924	35.2967	NM_144646	IGJ	NM_144646	7.853210073
16797403	431.224	55.1977	---	---	AY510104	7.812360162
16905600	215.332	27.6004	NM_133378	TTN	NM_133378	7.801776642
17011886	30.68319	3.94657	---	---	TCONS_000119	7.774657202
17052737	236.6025	30.8979	NM_002652	PIP	NM_002652	7.657559252
16797336	1011.111	137.856	---	---	BC065733	7.334534101
16709458	62.47979	8.59147	NM_198795	TDRD1	NM_198795	7.272303279
16927785	644.8411	89.7019	---	---	ENST00000385	7.188714126
16771434	345.3532	48.3531	---	---	---	7.142316182
16797479	208.7004	30.5188	---	---	AK289373	6.838423148

TABLE 1: Genes up-regulated in Stage 3 Prostate Cancer

16900096	903.503	132.689	BC032451	IGK@	BC032451	6.809198379
16797690	29.51886	4.3362	---	---	---	6.807538025
16983765	307.1956	46.5068	NM_000908	NPR3	NM_000908	6.60539246
17056833	492.772	75.8418	---	---	AF085845	6.497369457
16900132	989.4561	156.082	ENST0000047998	IGKV1-16	ENST00000479	6.339335093
16805899	84.82558	13.7971	NR_024387	CXADRP2	NR_024387	6.148064234
17049676	245.1839	40.3748	NM_000602	SERPINE1	NM_000602	6.072691823
16907831	255.7697	42.375	NM_079420	MYL1	NM_079420	6.035863127
16798951	232.3069	38.7016	NM_013372	GREM1	NM_013372	6.002520312
16797409	411.3531	68.8531	AK090461	IGHD	AK090461	5.974358453
16854408	119.953	20.2775	NR_024259	LOC728606	NR_024259	5.915574363
16889743	44.3531	7.52407				5.894826631
16900100	43.68627	7.48967	ENST0000049616	IGKV1-5 // I	ENST00000496	5.832872921
16800630	69.86362	12.0546	NM_032413	C15orf48	NM_032413	5.795588745
16650201	15.91893	2.76689	---	---	---	5.753365692
16927810	202.2847	35.167	ENST0000039031	IGLV2-18 //	ENST00000390	5.752120313
16927756	267.8531	47.3531	---	---	BC073769	5.656502537
17122982	1407.769	249.319	---	---	---	5.646459213
16882720	1287.047	229.739	ENST0000049857	IGKV1-39	ENST00000498	5.602204063
16797504	101.5519	18.2327	---	---	AK303185	5.569773615
17052973	371.3268	66.7751	NM_014141	CNTNAP2	NM_014141	5.5608556
16988423	120.4087	21.8845	NM_016644	PRR16	NM_016644	5.502005757
17124830	80.73445	14.9903	---	---	---	5.385786656
16771433	260.3532	48.8531				5.329307659
16859795	1084.581	204.641	NM_004864	GDF15	NM_004864	5.299922938
16962689	88.84472	16.8925	NM_001146686	GEMC1	NM_001146686	5.259414971
16771432	246.8531	47.3531				5.213028247
16882723	580.8939	111.433	ENST0000046517	IGKV1-37	ENST00000465	5.212929169
16725041	85.73141	16.4861	AK304258	FAM111B	AK304258	5.200220671
16826110	27.76014	5.36225	---	---	ENST00000450	5.176953572
16771448	82.3531	15.9189				5.173281119
17041111	572.3929	111.749	---	---	---	5.122120785
17031045	542.9544	106.942	---	---	---	5.077082761
17041030	576.8534	113.619	---	---	---	5.07708123
16908616	17.87779	3.52748				5.06815211
17007118	569.0654	112.403	NM_001002029	C4B	NM_001002029	5.062737817
17035749	576.0376	113.78	---	---	---	5.062736037
16797353	386.5006	76.4972	ENST0000039054	IGHG4 // IGH	ENST00000390	5.052479692
17068918	42.02176	8.31722	---	---	ENST00000520	5.052382305
17007048	554.1838	109.97	NM_007293	C4A	NM_007293	5.039431842
17028176	585.5333	116.728	---	---	---	5.016241246
16709245	114.3195	22.9462	NM_000681	ADRA2A	NM_000681	4.982060234
16927840	137.7116	27.8571	ENST0000039031	IGLV3-1	ENST00000390	4.943508037
17017217	95.35309	19.3674				4.923373416
16797563	130.4524	26.5211	BC089417 // BC	IGHG1 // IGH	BC089417	4.918817439
16652211	9.485942	1.96101	---	---	---	4.83727365
16797459	25.60906	5.2969	BC041037	IGHM	BC041037	4.834722321
16797603	72.77382	15.0733	AK127409	IGHG1	AK127409	4.828011212
16778241	574.0981	120.903	NM_006475	POSTN	NM_006475	4.748415053

TABLE 2: Genes down-regulated in Stage 3 Prostate Cancer

Probe ID	Stage 3	Normal	GeneAccession	Gene Symbol	mRNA Accession	Decreased Exp
16940033	11.826	638.68	---	---	ENST000004272!	0.018516648
17122144	6.2387	259.35	---	---	---	0.024054931
16656659	6.0239	215.79	---	---	---	0.027915161
16656655	3.2542	113.61	---	---	---	0.028642713
16656633	10.449	318.04	---	---	---	0.032854283
16656621	5.4577	138.86	---	---	---	0.039303895
16940012	250.17	5124	NM_003241	TGM4	NM_003241	0.048823623
16656649	12.12	243.19	---	---	---	0.049838728
16656663	12.839	242.21	---	---	---	0.053008222
16656629	29.143	521.81	---	---	---	0.055849803
16838249	22.82	391.11	---	---	TCONS_0002513	0.058346925
17018396	5.6622	94.853	---	---	---	0.059694448
16656641	40.009	669.33	---	---	---	0.059774216
16946280	24.948	415.38	NM_022131	CLSTN2	NM_022131	0.060061471
16656643	20.828	333.68	---	---	---	0.06241899
16775083	42.011	659.24	NM_006418	OLFM4	NM_006418	0.063725796
16656665	14.388	225.43	---	---	---	0.063825701
16873204	13.385	196.08	---	---	TCONS_l2_00013	0.068261563
16656647	49.361	712.8	---	---	---	0.06924903
16656657	33.037	466.29	---	---	---	0.070849745
16744602	16.991	236.93	NM_001077639	FAM55D	NM_001077639	0.071714189
16656623	25.087	346.64	---	---	---	0.072372787
16852849	40.26	550.43	NM_080475	SERPINB11	NM_080475	0.07314271
16877318	8.6936	114.85	---	---	TCONS_0000358	0.075693159
16887010	10.775	140.67	---	---	ENST0000041896	0.076600668
16656631	55.242	719.1	---	---	---	0.076821157
16656645	21.43	269.83	---	---	---	0.079420414
16656625	21.654	271.14	---	---	---	0.079861297
16656627	21.093	243.64	---	---	---	0.086574312
16781423	2.8489	30.936	---	---	---	0.092088406
16869919	22.381	241.27	---	---	TCONS_0002720	0.092762895
17095327	3.5275	37.353	---	---	---	0.094435991
16656637	100.16	1030.9	---	---	---	0.097158633
16656653	5.0109	48.755	---	---	---	0.102777474
16656365	4.0486	38.353	---	---	---	0.105560776
16681986	5.2349	47.836	---	---	AK055853	0.109432777
16688386	8.3121	73.184	NM_001114120	DEPDC1	NM_001114120	0.113578084
17020069	30.737	270.55	NM_145740	GSTA1	NM_145740	0.113609429
16656661	80.853	708.35	---	---	---	0.114142303
16873190	34.984	303.99	NM_001102597	CEACAM20	NM_001102597	0.11508526
16990124	4.8672	41.853	---	---	---	0.11629143
16809872	17.459	146.24	---	---	---	0.119387685
17118860	270.37	2236.1	---	---	---	0.120908221
17009219	30.692	251.43	---	---	---	0.122070117
16886806	8.3794	68.485	NM_001017920	DAPL1	NM_001017920	0.122354193
16656609	13.688	111.69	---	---	---	0.122545255
16938645	9.5682	73.627	---	---	---	0.129953918
16821186	11.575	89.039	NM_005752	CLEC3A	NM_005752	0.130000478

TABLE 2: Genes down-regulated in Stage 3 Prostate Cancer

16781648	13.8	104.2	NM_014579	SLC39A2	NM_014579	0.132431439
17009989	3.5174	25.379	---	---	---	0.138599132
16888229	32.909	237.35	---	---	---	0.138650689
16656639	64.075	446.44	---	---	---	0.143524513
16844735	23.823	165.57	NM_153490	KRT13	NM_153490	0.143881982
16977810	11.475	78.582	NM_178135	HSD17B13	NM_178135	0.146031581
16771533	4.4345	30.353	---	---	---	0.146096891
16789823	6.9965	47.319	---	---	BC015119	0.147857638
17020059	11.515	77.374	NM_000846	GSTA2	NM_000846	0.148820521
16650169	5.0524	33.713	---	---	---	0.149864505
17020258	31.145	204.57	NM_021073	BMP5	NM_021073	0.152249232
17084374	9.1675	59.874	---	---	---	0.1531148
16790138	14.381	93.802	---	---	AK026285	0.153312867
16735833	35.353	227.85	---	---	---	0.155157492
17083845	5.7656	36.864	---	---	ENST000005171!	0.156402214
16673840	70.694	419.46	---	---	---	0.168534858
16968307	136.15	806.33	NR_026555	GDEP	NR_026555	0.168856087
16656603	4.4345	25.853	---	---	---	0.1715249
16771531	4.8672	28.353	---	---	---	0.171661985
16744585	5.5649	32.339	NM_152315	FAM55A	NM_152315	0.172078762
16745561	162.85	932.45	---	---	---	0.17464698
16656607	43.983	251.08	---	---	---	0.175174535
16698502	37.838	211.67	---	---	---	0.178762811
16654487	6.8124	37.411	---	---	---	0.182095968
16656247	2.8242	15.495	---	---	---	0.182260341
16654391	4.1713	22.781	---	---	---	0.183103006
16805995	10.55	57.4	---	---	---	0.183794102
16906089	13.132	71.233	---	---	---	0.184349488
16656015	13.046	69.471	---	---	---	0.187792919
16652875	13.203	70.054	---	---	---	0.188474016
16656289	11.398	59.602	---	---	---	0.191227413
16655523	2.7914	14.564	---	---	---	0.191664653
16653061	11.394	58.497	---	---	---	0.194775024
16656619	21.373	108.99	---	---	---	0.196096012
16656651	6.8494	34.853	---	---	---	0.196521314
16859166	10.199	51.53	NM_023944	CYP4F12	NM_023944	0.197931852
16656153	6.9357	34.667	---	---	---	0.200064557
16651031	3.7924	18.865	---	---	---	0.201031031
16751999	12.581	62.353	---	---	---	0.201766071
17068615	9.4859	46.353	---	---	---	0.204645256
16771468	4.0486	19.367	---	---	---	0.209040745
16651193	9.1221	43.522	---	---	---	0.20959556
16693526	10.319	48.853	---	---	---	0.211223443
17092748	8.3227	39.388	---	---	---	0.211301023
17011664	32	151.31	---	---	---	0.211485799
17075920	17.939	83.887	---	---	---	0.213850672
16835637	7.5241	34.853	---	---	---	0.215879563
16790255	6.2286	28.353	---	---	---	0.219679295
16993425	20.361	91.853	---	---	---	0.221670726



TABLE 3

# MALIGNANT-SPECIFIC MBD1 PEAKS COMMON TO ALL SAMPLES

chromosome	start	stop
chr1	12523050	14268300
chr1	116128050	116608500
chr2	0	758700
chr2	6571500	6805800
chr2	91831350	91876950
chr2	228859650	229637550
chr3	60972750	61300200
chr3	61425150	61445850
chr3	90418200	90427350
chr3	196933800	197440200
chr3	197896650	198295500
chr5	0	56250
chr5	56550	400650
chr5	79544250	80367900
chr6	73826250	74213850
chr6	74214000	74715900
chr6	160291200	160698450
chr6	161017050	161046750
chr7	149883750	152072250
chr7	158249250	158423700
chr8	98590350	99118500
chr10	37540500	37678800
chr11	51373350	51568050
chr12	7459200	8020500
chr12	8021700	8246850
chr14	72603750	73275750
chr14	73275900	73314900
chr14	88099800	88284750
chr17	21554250	21902400
chr18	4336500	5065350
chr18	7039200	7672950
chrX	138940350	142020000

TABLE 4

# GENE-CONTAINING REGIONS OVERLAPPING MALIGNANT- SPECIFIC MBD1 PEAKS

chromosome	start	stop	decreased	increased
chr1	12523050	14268300		
chr1	116128050	116608500	SLC22A15	CASQ2
chr2	0	408600	FAM110C	
chr2	91831350	91832550		
chr2	228893250	229140900		
chr3	60972750	61300200		
chr3	196933800	197425800		
chr3	197426250	197440200		KIAA0226
chr3	197896650	197929950		
chr3	197931000	198295500		
chr5	56550	400650	PLEKHG4B, PDCD6	CCDC127
chr5	79947900	79999800		DHFR
chr6	74082600	74213850		
chr6	74214000	74517000		
chr6	160291200	160698450		
chr6	161017050	161046750		
chr7	150249900	150852300		
chr7	158249250	158322450		PTPRN2
chr7	158323800	158423700		PTPRN2
chr8	98606250	98729250		
chr8	98790900	99118500	HRSP12	SNORNA72
chr11	51374400	51471450		
chr11	51471750	51567300		
chr12	7459200	8020500		
chr12	8021700	8023200		
chr14	72966600	73275750		DPF3
chr14	73275900	73314900		DPF3
chr17	21554250	21902400		
chr18	7039200	7224750		
chrX	138940350	140755650		
chrX	140755950	140875800		



TABLE 5

## MALIGNANT-SPECIFIC MeCP2 PEAKS COMMON TO ALL SAMPLES

chromosome	start	stop
chr1	212202300	213467700
chr1	214479900	214782600
chr1	222022200	222201450
chr2	89222850	89830350
chr2	111501750	113411100
chr2	133046400	133685700
chr2	158956800	159025200
chr2	228379500	228858900
chr3	59997150	60304200
chr3	75700650	75715950
chr3	140437800	140608050
chr4	44100	59250
chr4	49521150	49527900
chr4	111155100	111810300
chr4	185137350	186126900
chr5	151869450	151945800
chr5	153061200	154112250
chr7	152664300	153193650
chr8	43097100	43235700
chr9	66837000	66840000
chr9	126298350	130003950
chr10	106217550	106511400
chr10	129684750	130794900
chr10	131132550	132382200
chr11	47570250	48229800
chr11	123208500	124324350
chr11	128769000	129463800
chr11	132696150	133383600
chr15	93206700	93763500
chr16	65125350	65733300
chr18	7039200	7672950
chr20	17567250	19164000
chr20	23884350	24613500
chr20	24613950	25904550
chr20	40813950	41133600
chr20	44077350	45155700
chr22	20658000	20689050
chrY	9992550	10017600

TABLE 6

# GENE-CONTAINING REGIONS OVERLAPPING MALIGNANT- SPECIFIC MeCP2 PEAKS

chromosome	start	end	decreased	increased
chr1	213434550	213467700		RPS6KC1
chr1	214479900	214782600		
chr2	89222850	89830350		
chr2	111501750	112198050	ACOXL	
chr2	133046400	133685700		LYPD1, NCKAP5
chr2	158990400	159025200		
chr2	228379500	228858900		
chr3	59997150	60107100		
chr3	75702900	75715650		
chr4	111155100	111810300		
chr4	185639100	186126900		
chr5	153061200	153146700	GRIA1	
chr8	43097400	43235700		
chr9	127475250	128050800		
chr10	106217550	106511400	SORCS3	
chr11	47834250	48229800		
chr11	123208500	124041300		
chr11	124041450	124324350		
chr11	128769000	129178350	C11orf45	
chr11	133287300	133383600		
chr15	93732150	93763500		
chr16	65125350	65692350		CDH11
chr18	7224900	7672950		PTPRM
chr20	17871000	18138750		
chr20	24493950	24613500		
chr20	24613950	25222500		
chr20	25223250	25423950	GIN51	
chr20	40958250	41133600		PTPRT
chr20	44077350	44654700		

**TABLE 7**

**MALIGNANT-SPECIFIC MBD4 PEAKS**

**COMMON TO THE TWO PATIENT SAMPLES**

Patient 3			Patient 6		
chr1	48451650	48705750	chr1	47541600	48588300
chr1	48451650	48705750	chr1	48685350	48703350
chr1	53847900	54832200	chr1	54372150	55172400
chr1	54833400	54921300	chr1	54372150	55172400
chr1	90298800	90887550	chr1	90763800	90771300
chr1	155525550	155858550	chr1	154907700	156195150
chr1	155858700	156243150	chr1	154907700	156195150
chr1	160212600	160679550	chr1	160212600	160447500
chr1	227740800	228058350	chr1	227613900	228042900
chr1	227740800	228058350	chr1	228043350	228110700
chr1	228058500	228422700	chr1	228216150	229731900
chr1	228058500	228422700	chr1	228043350	228110700
chr1	228423750	228610050	chr1	228216150	229731900
chr1	228612600	228870600	chr1	228216150	229731900
chr1	228870750	229166700	chr1	228216150	229731900
chr1	231021300	231344700	chr1	231122100	231463650
chr1	231021300	231344700	chr1	230768100	231121950
chr1	244950600	244960350	chr1	244468200	245469900
chr10	47668350	49775550	chr10	47629050	47941500
chr10	73386900	73392900	chr10	72137850	73530600
chr10	119401950	119546550	chr10	119192850	120390750
chr10	120360600	120952350	chr10	119192850	120390750
chr10	132578400	133241850	chr10	133174200	133264950
chr11	32156700	32287950	chr11	31635000	32223750
chr11	43654050	44115450	chr11	43877700	44000550
chr11	48868650	48876600	chr11	48868800	48877500
chr11	71920200	72774300	chr11	72122250	72505050
chr11	71920200	72774300	chr11	72505200	73249050
chr11	72776400	73114200	chr11	72505200	73249050
chr11	73114800	73189050	chr11	72505200	73249050
chr11	78640500	79104900	chr11	78370500	79122000
chr11	126216600	126509400	chr11	126349950	126376350
chr11	126216600	126509400	chr11	126092850	126349500
chr11	134626650	135086550	chr11	134589900	135086550
chr12	1986900	2656500	chr12	2205750	2294250
chr12	1986900	2656500	chr12	1696800	2205600
chr12	1986900	2656500	chr12	2294400	2343900
chr12	1986900	2656500	chr12	2344050	2356050
chr12	30933150	31372800	chr12	31322400	31562700
chr12	61923150	62091450	chr12	61856550	61927650
chr12	64015650	64087500	chr12	63922350	64744350
chr12	64860450	65359200	chr12	64878750	65446350
chr13	24174000	24556800	chr13	23889900	24759750

chr13	74046750	74453700	chr13	73883700	74180850
chr14	73446600	73521900	chr14	73098450	74237700
chr14	73522200	73546650	chr14	73098450	74237700
chr14	78194250	78312600	chr14	77162850	78426150
chr14	78312750	78394050	chr14	77162850	78426150
chr14	100137000	100677750	chr14	99734550	100515900
chr15	28277100	28783650	chr15	28388550	28462050
chr15	72331200	73209600	chr15	72578850	72604050
chr15	72331200	73209600	chr15	73065450	73180350
chr15	93207600	93427800	chr15	93285750	93696600
chr15	93747600	94232400	chr15	93699600	93843300
chr16	27456300	27687900	chr16	26827500	27614700
chr16	27688050	28160550	chr16	28032300	28137000
chr16	54940800	55819350	chr16	55255050	55279500
chr17	26460300	26694900	chr17	26480550	26912250
chr17	26695050	27083550	chr17	26480550	26912250
chr17	26695050	27083550	chr17	26912850	27257100
chr17	27084900	27775950	chr17	26912850	27257100
chr17	27084900	27775950	chr17	27257400	27544350
chr17	27084900	27775950	chr17	27544500	28207350
chr2	9548700	9585750	chr2	9362250	9869550
chr2	9586800	10096650	chr2	9362250	9869550
chr2	9586800	10096650	chr2	9869700	10677000
chr2	10096800	11310000	chr2	9869700	10677000
chr2	26814300	27428550	chr2	26421600	27201750
chr2	26814300	27428550	chr2	27201900	27342900
chr2	42857400	42873750	chr2	42378000	43356000
chr2	72019200	72273000	chr2	72181500	72191700
chr2	72019200	72273000	chr2	71745300	72181350
chr2	72019200	72273000	chr2	72192000	72203850
chr2	109627200	109678050	chr2	109185900	109815750
chr2	109678950	109815750	chr2	109185900	109815750
chr2	120928800	121146150	chr2	120649200	121886550
chr2	124640550	124660350	chr2	124621950	124650000
chr2	127222800	128156850	chr2	127952100	129696300
chr2	128157450	128969100	chr2	127952100	129696300
chr2	128970000	128989950	chr2	127952100	129696300
chr2	128990100	129090000	chr2	127952100	129696300
chr2	130040550	130963500	chr2	129964200	130171950
chr2	132988800	132994950	chr2	132988050	132995850
chr2	178087950	178348200	chr2	178223400	178562100
chr2	178087950	178348200	chr2	178053300	178223250
chr2	225591600	225905850	chr2	225636450	225993000
chr2	240399750	240759600	chr2	240016650	240600450
chr20	7518750	7867200	chr20	7718550	7800750
chr20	24716550	25172100	chr20	25063500	25269750
chr20	25174950	25584150	chr20	25063500	25269750
chr20	29652450	29804100	chr20	29652000	29804400
chr21	9703500	9710550	chr21	9703200	9712050
chr21	10766850	10770450	chr21	10765050	10772850
chr22	27620250	28537200	chr22	27768150	28350300
chr3	13620300	13929300	chr3	13477200	13768800

chr3	13929600	14577900	chr3	13961250	14026650
chr3	32480700	33645900	chr3	33311250	33322650
chr3	54449850	54937350	chr3	54366000	54661500
chr3	90475350	90486000	chr3	90478200	90486300
chr3	150364050	150446550	chr3	150158550	150643350
chr3	185415450	186051450	chr3	185187000	185874750
chr4	12287100	12671400	chr4	12613800	12641850
chr4	12671550	13084050	chr4	12816600	12863250
chr4	12671550	13084050	chr4	12960900	13053300
chr4	53306250	53427600	chr4	53088600	53469150
chr4	53428200	53449650	chr4	53088600	53469150
chr4	68266650	68326500	chr4	68266950	68311950
chr4	166484100	166505850	chr4	166327500	166925550
chr4	185763000	186126900	chr4	185818200	185949600
chr5	7547400	7684350	chr5	7538700	7667550
chr5	7547400	7684350	chr5	7667700	7868550
chr5	7843200	7954950	chr5	7667700	7868550
chr5	17530500	17580600	chr5	17526600	17584650
chr5	149702850	150479100	chr5	150476100	150828300
chr5	149702850	150479100	chr5	150170100	150256800
chr5	150795000	150810750	chr5	150476100	150828300
chr5	170943900	171013650	chr5	170553900	171279150
chr5	171113400	171366150	chr5	170553900	171279150
chr5	171113400	171366150	chr5	171279450	171646950
chr6	1698000	1851750	chr6	1765800	3367650
chr6	2897550	3328650	chr6	1765800	3367650
chr6	3328800	3411900	chr6	1765800	3367650
chr6	121773900	121803900	chr6	121771500	121842150
chr7	6247650	6435000	chr7	6346350	6384300
chr7	28077150	28401300	chr7	28136100	28167150
chr7	58022100	58028550	chr7	58023000	58032750
chr7	128146200	128715900	chr7	128673300	128678400
chr7	135332250	135603150	chr7	134878200	135483450
chr7	154494150	155272500	chr7	154607100	154612950
chr8	0	758100	chr8	482250	837600
chr8	758400	791550	chr8	482250	837600
chr8	20871600	20932800	chr8	20748900	20976150
chr8	40601550	41733000	chr8	41583150	42248400
chr8	86576400	86726550	chr8	86574300	86728050
chr8	101609700	101965950	chr8	101818800	101932800
chr8	101609700	101965950	chr8	101932950	101940600
chr8	124796850	124904100	chr8	124601100	125125650
chr8	143829600	144751500	chr8	143756250	144110700
chr9	86948700	87136200	chr9	87029850	87072750
chr9	99577950	99771450	chr9	99407850	100405650
chrX	14254650	15004200	chrX	14191950	14296650
chrX	111008250	111412950	chrX	111210900	111227550
chrY	9966450	9972000	chrY	9959550	9973350

TABLE 8: MBD1-SPECIFIC PEAKS IN  
SAMPLE #3 NORMAL TISSUE

Chromos Location	Start	Stop	Interval Length	Maximum Extended Reads in Window	Reads Per Million (RPM)	Mann- Whitney p-value	015_Input: total reads in region	012: total reads in region	p-value (Sample ID vs. 015_Input)	scaled fold change (Sample ID vs. 015_Input)
1	143280363	143280632	269	29	24.6	0.000949	7	29	1.15E-05	4.5
1	152071255	152071455	200	28	23.8	2.22E-07	0	28	1.59E-10	29.0
2	89855492	89855725	233	26	22.1	0.007076	5	30	6.12E-07	6.2
2	105451119	105451403	284	42	35.7	0.018127	0	42	2.01E-15	43.0
2	116630336	116630628	292	37	31.4	0.00866	0	37	1.13E-13	38.0
2	124311717	124311917	200	25	21.2	8.85E-05	0	25	1.79E-09	26.0
2	171046320	171046523	203	43	36.5	4.68E-06	0	42	2.01E-15	43.0
3	132593896	132594115	219	29	24.6	0.019775	0	28	1.59E-10	29.0
4	49099603	49099937	334	59	50.1	0.000247	29	65	1.44E-06	2.7
4	78671563	78671765	202	44	37.4	0.01491	0	44	4.02E-16	45.0
5	90368423	90368623	200	26	22.1	3.16E-05	0	26	7.98E-10	27.0
5	129033693	129033894	201	59	50.1	6.18E-08	0	59	2.26E-21	60.0
6	57420994	57421194	200	26	22.1	0.005279	0	26	7.98E-10	27.0
6	63742069	63742269	200	35	29.7	1.81E-08	0	33	2.84E-12	34.0
7	81484848	81485147	299	29	24.6	0.020106	0	29	7.12E-11	30.0
7	100243496	100243697	201	31	26.3	0.0009	0	32	6.35E-12	33.0
7	133793925	133794126	201	31	26.3	0.005249	0	29	7.12E-11	30.0
8	42274229	42274429	200	33	28.0	1.00E-04	0	37	1.13E-13	38.0
8	60439838	60440045	207	35	29.7	1.28E-07	0	35	5.66E-13	36.0
8	67782070	67782464	394	31	26.3	0.00302	2	31	2.42E-09	12.2
10	8536156	8536492	336	35	29.7	6.73E-09	0	35	5.66E-13	36.0
10	19049775	19049975	200	34	28.9	0.015793	0	34	1.27E-12	35.0
11	36507360	36507561	201	27	22.9	0.000911	0	22	2.00E-08	23.0
11	48961677	48962064	387	42	35.7	0.000786	1	43	2.23E-14	24.3
12	34841440	34841733	293	36	30.6	0.019228	24	36	0.0121762	1.8
12	66148413	66148613	200	39	33.1	9.48E-10	0	39	2.26E-14	40.0
12	73995168	73995418	250	34	28.9	0.004634	0	34	1.27E-12	35.0
13	85554050	85554250	200	35	29.7	6.63E-09	0	35	5.66E-13	36.0
13	96096292	96096492	200	27	22.9	3.81E-07	0	27	3.57E-10	28.0
14	88379624	88379824	200	38	32.3	0.01164	2	39	5.90E-12	15.3
16	22367793	22367993	200	35	29.7	6.63E-09	1	35	1.15E-11	19.9
16	33963881	33964262	381	58	49.3	0.000724	19	70	6.01E-11	4.3
16	46400199	46400519	320	33	28.0	0.004751	29	42	0.00996132	1.8
18	50002111	50002311	200	28	23.8	4.30E-07	0	27	3.57E-10	28.0
19	9710522	9710725	203	27	22.9	0.004014	20	27	0.0536516	1.6
19	37333743	37333943	200	26	22.1	3.61E-07	0	27	3.57E-10	28.0
20	11412592	11412792	200	26	22.1	0.020913	0	26	7.98E-10	27.0
21	10856733	10857069	336	41	34.8	0.003161	9	43	1.91E-08	5.3
22	18568833	18569033	200	25	21.2	0.018258	0	25	1.79E-09	26.0
Y	13808692	13808892	200	25	21.2	0.006342	21	25	0.121031	1.4
Y	58853928	58854202	274	29	24.6	0.020345	25	32	0.0543748	1.6
Y	58886170	58886424	254	28	23.8	0.008838	18	28	0.0199038	1.9

**TABLE 9: GENES NEAREST MBD1-SPECIFIC PEAKS IN  
SAMPLE #3 NORMAL TISSUE**

Chromos Location	Transcript Start	Transcript Stop	Strand	Transcript ID	Gene Symbol	Distance to TSS	Percent overlap with gene	Start	Stop
1	152056620	152061541	-	NM_001008536	TCHHL1	-9714	0.000	152071255	152071455
2	105421883	105467935	-	NR_037883	AC018730.1	16532	0.619	105451119	105451403
2	171034655	171511675	+	NM_001083615	MYO3B	11665	0.043	171046320	171046523
2	171034655	171511675	+	NM_138995	MYO3B	11665	0.043	171046320	171046523
2	171034655	171511675	+	NR_045682	MYO3B	11665	0.043	171046320	171046523
2	171034655	171511675	+	NR_045683	MYO3B	11665	0.043	171046320	171046523
2	171034655	171511675	+	NR_045684	MYO3B	11665	0.043	171046320	171046523
4	78634541	78740545	-	NM_001286790	CNOT6L	68780	0.192	78671563	78671765
4	78634541	78740545	-	NM_144571	CNOT6L	68780	0.192	78671563	78671765
5	89854617	90460034	+	NM_032119	GPR98	513806	0.033	90368423	90368623
5	89854617	90460034	+	NR_003149	GPR98	513806	0.033	90368423	90368623
5	128796103	129074377	+	NM_133638	ADAMTS19	237590	0.073	129033693	129033894
6	57182415	57513377	+	NM_000947	PRIM2	238579	0.061	57420994	57421194
7	100218039	100239174	-	NM_003227	TFR2	-4322	0.000	100243496	100243697
7	100240726	100254085	-	NM_016188	ACTL6B	10388	1.512	100243496	100243697
8	42273980	42358979	-	NM_001257180	SLC20A2	84550	0.236	42274229	42274429
8	42273980	42396656	-	NM_001257181	SLC20A2	122227	0.164	42274229	42274429
8	42273980	42397357	-	NM_006749	SLC20A2	122928	0.163	42274229	42274429
8	67782984	67814015	+	NM_001136161	MCMDC2	-520	0.000	67782070	67782464
8	67782984	67834284	+	NM_173518	MCMDC2	-520	0.000	67782070	67782464
8	67783737	67817600	+	NM_001136160	MCMDC2	-1273	0.000	67782070	67782464
11	36505317	36531864	-	NM_004620	TRAF6	24303	0.761	36507360	36507561
11	36505317	36531864	-	NM_145803	TRAF6	24303	0.761	36507360	36507561
13	96085853	96232011	+	NM_001160100	CLDN10	10439	0.138	96096292	96096492
13	96085853	96232011	+	NM_182848	CLDN10	10439	0.138	96096292	96096492
16	22357257	22385939	-	NM_001802	CDR2	17946	0.701	22367793	22367993
16	33961052	33962504	-	NR_038368	LINC00273	-1377	0.000	33963881	33964262
18	49866542	51062274	+	NM_005215	DCC	135569	0.017	50002111	50002311
19	37309224	37341690	-	NM_001242800	ZNF790	7747	0.619	37333743	37333943
19	37309224	37341170	-	NM_001242801	ZNF790	7227	0.629	37333743	37333943
19	37309224	37328930	-	NM_001242802	ZNF790	-4813	0.000	37333743	37333943
19	37309224	37329287	-	NM_206894	ZNF790	-4456	0.000	37333743	37333943
19	37341260	37370478	+	NM_003419	ZNF345	-7317	0.000	37333743	37333943
19	37341778	37370478	+	NM_001242472	ZNF345	-7835	0.000	37333743	37333943
19	37341778	37370478	+	NM_001242474	ZNF345	-7835	0.000	37333743	37333943
19	37341788	37384121	+	NR_038362	ZNF345	-7845	0.000	37333743	37333943
19	37342553	37370478	+	NM_001242475	ZNF345	-8610	0.000	37333743	37333943
19	37342553	37370478	+	NM_001242476	ZNF345	-8610	0.000	37333743	37333943
22	18560686	18573798	+	NM_001199319	PEX26	8147	1.533	18568833	18569033
22	18560686	18573798	+	NM_017929	PEX26	8147	1.533	18568833	18569033
22	18560760	18573798	+	NM_001127649	PEX26	8073	1.542	18568833	18569033

TABLE 10: GENES NEAREST MBD1-SPECIFIC  
PEAKS IN SAMPLE #4 NORMAL TISSUE

chromos location	transcript start	transcript stop	strand	Transcript ID	Gene Symbol	Distance to TSS	Percent overlap with gene	Start	Stop
1	169101768	169337202	-	NM_013330	NME7	221772	0.0854	169115230	169115430
1	169101768	169337202	-	NM_197972	NME7	221772	0.0854	169115230	169115430
1	169101768	169337202	-	NR_104229	NME7	221772	0.0854	169115230	169115430
2	140988996	142889271	-	NM_018557	LRP1B	220073	0.0106	142668997	142669198
5	57878939	58147407	+	NM_138453	RAB3C	221851	0.1382	58100790	58101160
5	72848025	72861512	-	NM_023039	ANKRA2	0	0.8749	72861395	72861643
5	72861566	72879203	+	NM_001284430	UTP15	0	0.4422	72861395	72861643
5	72861566	72879203	+	NM_001284431	UTP15	0	0.4422	72861395	72861643
5	72861566	72879203	+	NM_032175	UTP15	0	0.4422	72861395	72861643
6	121400640	121655647	-	NM_152730	TBC1D32	132117	0.0980	121523281	121523530
6	121400640	121655647	-	NR_104452	TBC1D32	132117	0.0980	121523281	121523530
10	131265454	131565784	+	NM_002412	MGMT	152400	0.0673	131417854	131418055
14	89290978	89344336	+	NM_144596	TTC8	49741	0.3767	89340719	89340919
14	89290978	89344336	+	NM_198309	TTC8	49741	0.3767	89340719	89340919
14	89290978	89344336	+	NM_198310	TTC8	49741	0.3767	89340719	89340919
15	50150435	50411420	-	NM_024837	ATP8B4	190400	0.0770	50220820	50221020
15	50150435	50411420	-	NR_073596	ATP8B4	190400	0.0770	50220820	50221020
15	50150435	50411420	-	NR_073597	ATP8B4	190400	0.0770	50220820	50221020
15	50150435	50405514	-	NR_073598	ATP8B4	184494	0.0788	50220820	50221020
22	25423941	25593416	+	NM_001145206	KIAA1671	65827	0.1227	25489768	25489975



TABLE 11: GENES NEAREST MBD1-  
SPECIFIC PEAKS IN SAMPLE #6 NORMAL

chromos location	transcript start	transcript stop	strand	Transcript ID	Gene Symbol	Distance to TSS	Percent overlap with gene	Start	Stop
2	101623690	101767847	-	NM_001102426	TBC1D8	61030	0.2178	101706504	101706817
2	232826293	233208679	+	NM_001257281	DIS3L2	192177	0.0552	233018470	233018680
2	232826293	233201909	+	NM_152383	DIS3L2	192177	0.0562	233018470	233018680
2	232826293	233201909	+	NR_046476	DIS3L2	192177	0.0562	233018470	233018680
2	232826293	233201909	+	NR_046477	DIS3L2	192177	0.0562	233018470	233018680
6	169857303	170102160	-	NM_001202550	WDR27	213321	0.0829	169888637	169888839
6	169857303	170102160	-	NM_182552	WDR27	213321	0.0829	169888637	169888839
9	138898383	138987132	-	NM_144653	NACC2	19850	0.2310	138967078	138967282
15	25923860	26108350	-	NM_024490	ATP10A	26610	0.1100	26081538	26081740
16	9847265	10276612	-	NM_000833	GRIN2A	296676	0.0468	9979736	9979936
16	9847265	10276264	-	NM_001134407	GRIN2A	296328	0.0469	9979736	9979936
16	9856532	10275925	-	NM_001134408	GRIN2A	295989	0.0479	9979736	9979936

TABLE 12: MBD1-SPECIFIC PEAKS IN  
SAMPLE #3 MALIGNANT TISSUE

Chromos Location	Start	Stop	Interval Length	Maximum Extended Reads in Window	Reads Per Million (RPM)	Mann- Whitney p-value	014_ total reads in region	016_Input : total reads in region	p-value (Sample ID vs. 016_Input)	scaled fold change (Sample ID vs. 016_Input)
1	121357889	121358132	243	27	22.5	0.0059903	31	23	0.170312	1.3
1	149990803	149991003	200	25	20.8	0.0005322	24	0	5.95E-08	25.0
2	41890203	41890403	200	36	30.0	3.76E-05	34	1	1.05E-09	17.5
2	119691328	119691528	200	26	21.7	1.00E-06	25	0	2.97E-08	26.0
2	148129583	148129869	286	29	24.1	0.0015445	29	0	1.86E-09	30.0
3	38759373	38759739	366	29	24.1	1.31E-07	29	0	1.86E-09	30.0
3	137367585	137367928	343	38	31.6	5.10E-07	38	3	5.23E-09	9.8
4	167020612	167020912	300	77	64.1	6.66E-17	77	0	6.58E-24	78.0
4	174676149	174676349	200	25	20.8	1.00E-06	25	0	2.97E-08	26.0
4	183354639	183354839	200	35	29.1	0.0045362	30	0	9.29E-10	31.0
5	11550571	11550772	201	47	39.1	0.0172161	47	0	7.08E-15	48.0
6	61914121	61914321	200	25	20.8	0.0016718	30	24	0.248142	1.2
6	87549450	87549843	393	28	23.3	0.00362	28	0	3.72E-09	29.0
7	28401093	28401293	200	38	31.6	4.06E-09	36	11	0.0001727	3.1
7	64868802	64869021	219	31	25.8	0.0039565	31	1	7.67E-09	16.0
7	79816175	79816546	371	35	29.1	6.63E-09	35	0	2.90E-11	36.0
8	52023643	52024022	379	33	27.5	1.81E-08	33	0	1.16E-10	34.0
8	100207668	100207868	200	27	22.5	6.07E-07	27	0	7.44E-09	28.0
8	121553804	121554029	225	48	40.0	2.49E-08	48	0	3.54E-15	49.0
8	121556941	121557141	200	66	55.0	6.61E-09	67	0	6.74E-21	68.0
9	66822750	66822973	223	28	23.3	0.0144307	28	21	0.195663	1.3
10	42375378	42375578	200	25	20.8	0.001981	28	21	0.195663	1.3
11	48587371	48587571	200	29	24.1	9.19E-07	29	0	1.86E-09	30.0
11	90514294	90514651	357	31	25.8	0.0143816	31	0	4.65E-10	32.0
14	105854291	105854491	200	31	25.8	2.17E-07	28	0	3.72E-09	29.0
16	33955206	33955451	245	34	28.3	0.0119406	40	13	0.000134	2.9
16	33957515	33958134	619	49	40.8	0.021599	96	30	1.52E-09	3.1
16	46427489	46427845	356	31	25.8	0.0148171	37	24	0.0617977	1.5
18	22882568	22882768	200	27	22.5	0.0068005	25	0	2.97E-08	26.0
18	26119072	26119285	213	31	25.8	8.33E-08	30	0	9.29E-10	31.0
21	10764032	10764269	237	32	26.6	0.0120824	33	6	7.14E-06	4.9
X	131584340	131584546	206	30	25.0	0.0001078	30	0	9.29E-10	31.0
Y	58884586	58884874	288	28	23.3	0.0055728	28	25	0.391723	1.1

**TABLE 14: GENES NEAREST MBD1-  
SPECIFIC PEAKS IN SAMPLE #4 MALIGNANT  
TISSUE**

chromos location	transcript start	transcript stop	strand	Transcript ID	Gene Symbol	Distance to TSS	Percent overlap with gene	Start	Stop
2	132905164	133015543	-	NR_027020	ANKRD30BL	-7193	0.0000	133022736	133022947
2	133014539	133014654	-	NR_031608	MIR663B	-8082	0.0000	133022736	133022947
2	179390717	179672151	-	NM_001256850	TTN	32106	0.1155	179639721	179640045
2	179390717	179672151	-	NM_001267550	TTN	32106	0.1155	179639721	179640045
2	179390718	179672151	-	NM_003319	TTN	32106	0.1155	179639721	179640045
2	179390718	179672151	-	NM_133378	TTN	32106	0.1155	179639721	179640045
2	179390718	179672151	-	NM_133432	TTN	32106	0.1155	179639721	179640045
2	179390718	179672151	-	NM_133437	TTN	32106	0.1155	179639721	179640045
2	179609136	179672151	-	NM_133379	TTN	32106	0.5157	179639721	179640045
4	57897237	57976552	-	NM_001553	IGFBP7	72412	0.2547	57903939	57904140
4	57898404	57976552	-	NM_001253835	IGFBP7	72412	0.2585	57903939	57904140
4	147628179	147867035	-	NM_031956	TTC29	50711	0.1461	147815976	147816324
5	179233388	179265078	+	NM_001142298	SQSTM1	18127	0.6342	179251515	179251715
5	179234003	179265078	+	NM_001142299	SQSTM1	17512	0.6468	179251515	179251715
5	179247842	179265078	+	NM_003900	SQSTM1	3673	1.1661	179251515	179251715
6	13363587	13486416	-	NM_001242628	GFOD1	75357	0.1726	13410848	13411059
6	13363587	13408370	-	NM_001242630	GFOD1	-2478	0.0000	13410848	13411059
6	13363587	13487870	-	NM_018988	GFOD1	76811	0.1706	13410848	13411059
6	131160488	131384463	-	NM_001135554	EPB41L2	34314	0.1152	131349892	131350149
6	131160488	131384463	-	NM_001135555	EPB41L2	34314	0.1152	131349892	131350149
6	131160488	131384463	-	NM_001199388	EPB41L2	34314	0.1152	131349892	131350149
6	131160488	131384463	-	NM_001431	EPB41L2	34314	0.1152	131349892	131350149
7	29234121	29553945	+	NM_004067	CHN2	206472	0.0688	29440593	29440812
8	135812763	135812851	-	NR_029666	MIR30B	-7678	0.0000	135820529	135820907
8	135817119	135817189	-	NR_029599	MIR30D	-3340	0.0000	135820529	135820907
9	21277687	21278563	-	NR_036676	IFNA22P	-9365	0.0000	21287928	21288322
9	119187504	119449543	-	NM_001184734	ASTN2	105835	0.0767	119343508	119343708
9	119187504	120177318	-	NM_014010	ASTN2	833610	0.0203	119343508	119343708
9	119187504	119449495	-	NM_198186	ASTN2	105787	0.0767	119343508	119343708
9	119187504	119449543	-	NM_198187	ASTN2	105835	0.0767	119343508	119343708
9	119187504	119449543	-	NM_198188	ASTN2	105835	0.0767	119343508	119343708
10	76586171	76792381	+	NM_001256468	KAT6B	202590	0.0975	76788761	76788961
10	76586171	76792381	+	NM_001256469	KAT6B	202590	0.0975	76788761	76788961
10	76586171	76792381	+	NM_012330	KAT6B	202590	0.0975	76788761	76788961
12	44229664	44783546	+	NM_001286211	TMEM117	321060	0.0462	44550724	44550979
12	44229664	44783546	+	NM_001286212	TMEM117	321060	0.0462	44550724	44550979
12	44229664	44783546	+	NM_001286213	TMEM117	321060	0.0462	44550724	44550979
12	44229664	44783546	+	NM_032256	TMEM117	321060	0.0462	44550724	44550979
12	120941082	120966965	-	NM_032314	COQ5	19128	0.7765	120947637	120947837
14	55034330	55260034	+	NM_001161576	SAMD4A	166374	0.1015	55200704	55200932
14	55034330	55260034	+	NM_015589	SAMD4A	166374	0.1015	55200704	55200932
15	52599480	52821248	-	NM_000259	MYO5A	4142	0.1429	52816790	52817106
15	52599480	52821248	-	NM_001142495	MYO5A	4142	0.1429	52816790	52817106
15	53805938	54051860	-	NM_182758	WDR72	110839	0.0817	53940821	53941021
15	53805938	54055076	-	NR_102334	WDR72	114055	0.0807	53940821	53941021
16	33961052	33962504	-	NR_038368	LINC00273	1541	0.0000	33960621	33960963
17	20483037	20484225	+	NM_001190790	CDRT15L2	2679	0.0000	20485716	20485916
17	30819628	31203903	-	NM_015194	MYO1D	362590	0.0523	30841113	30841313
17	79762010	79771890	+	NM_000160	GCCR	5058	2.1354	79767068	79767278
20	2082528	2129199	+	NM_080836	STK35	18104	0.5806	2100632	2100902
X	85403455	86087606	+	NM_001139514	DACH2	288284	0.0294	85691739	85691939
X	85403455	86087606	+	NM_053281	DACH2	288284	0.0294	85691739	85691939
X	85517971	86087606	+	NM_001139515	DACH2	173768	0.0353	85691739	85691939

TABLE 15: GENES NEAREST MBD1-SPECIFIC  
PEAKS IN SAMPLE #6 MALIGNANT TISSUE

chromos location	transcript start	transcript stop	strand	Transcript ID	Gene Symbol	Distance to TSS	Percent overlap with gene	Start	Stop
2	132905164	133015543	-	NR_027020	ANKRD30BL	-8377	0.0000	133023920	133024279
2	133014539	133014654	-	NR_031608	MIR663B	-9266	0.0000	133023920	133024279
7	156230483	156238283	-	NR_038232	LOC285889	-4792	0.0000	156243075	156243275
13	96743093	97491817	+	NM_153456	HS6ST3	674969	0.0279	97418062	97418270
14	93799565	94173690	+	NM_020818	UNC79	75169	0.0540	93874734	93874935
17	9153788	9479276	-	NM_004853	STX8	27909	0.0621	9451166	9451367
17	9153788	9479276	-	NR_033656	STX8	27909	0.0621	9451166	9451367
18	109065	122223	+	NR_033770	ROCK1P1	1624	3.7237	110689	111178
18	112256	112340	-	NR_107045	MIR8078	1162	0.0000	110689	111178
18	29843484	30050448	-	NM_001242409	GAREM	51507	0.1227	29998688	29998941
18	29843484	30050448	-	NM_022751	GAREM	51507	0.1227	29998688	29998941
21	14410487	14490572	+	NR_026916	ANKRD30BP2	-2013	0.0000	14408274	14408474